Genomics, Research and Healthcare

Sarion Bowers
Research Policy Advisor

wellcome trust sanger institute

Mission


Image: Magnus Manske

We use genome sequences to advance understanding of the biology of humans and pathogens in order to improve human health

wellcome trust sanger institute

Sanger profile

- At the centre of a collaborative network of science
- Large-scale science underpinned by high-throughput data production platforms and large IT infrastructure
- Pushing scientific boundaries
- Leadership in internationally collaborative research enterprises
- Consideration of impact on human health globally
- Data organisation, presentation and open release

wellcome trust
sanger
institute

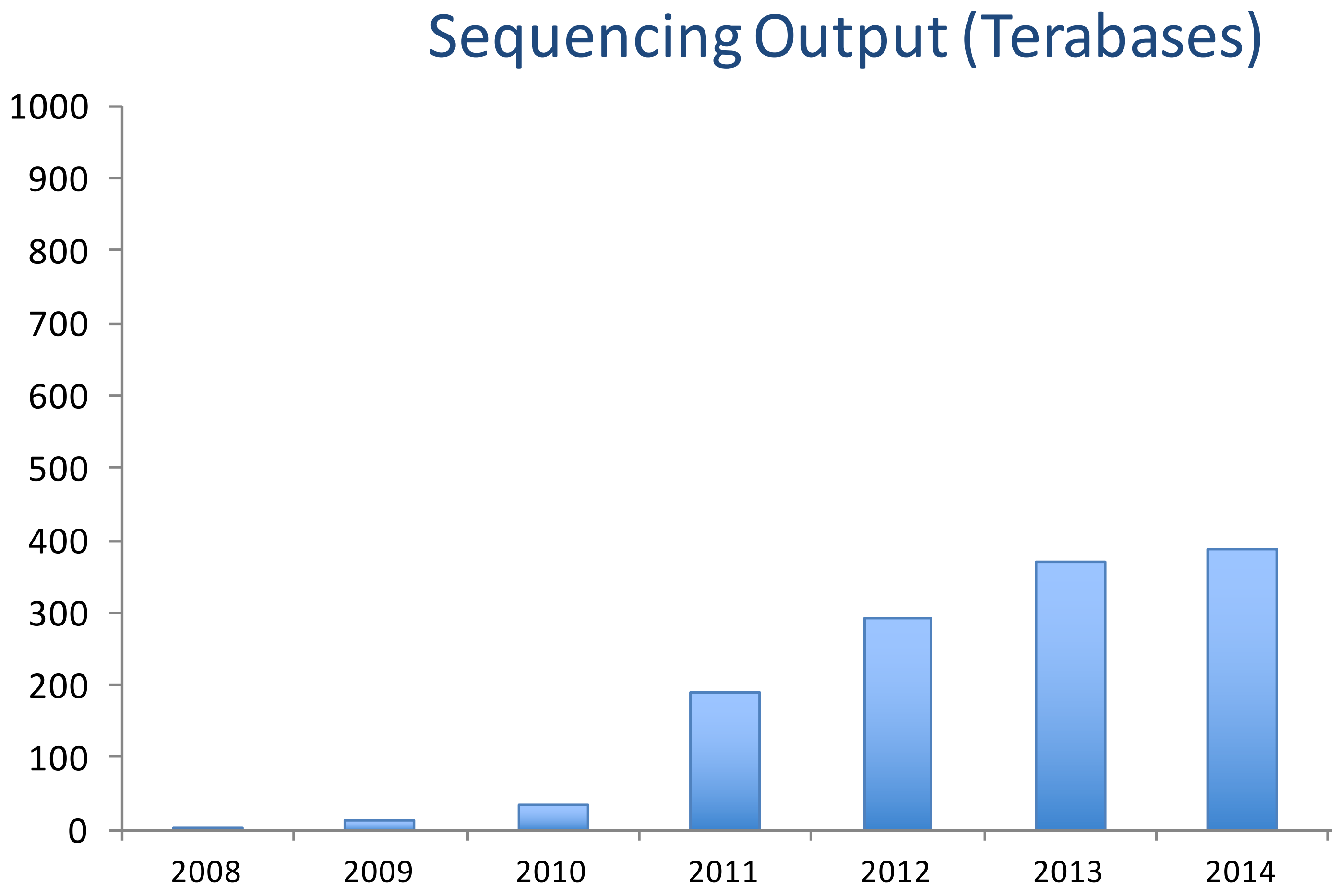
Scientific Operations

DNA pipelines

	Machines
Pacific Biosciences RS II	1
MiSeq	5
HiSeq 2000 V3	17
HiSeq 2500 V4	10
XTen system	10
Total	43

Scientific Operations

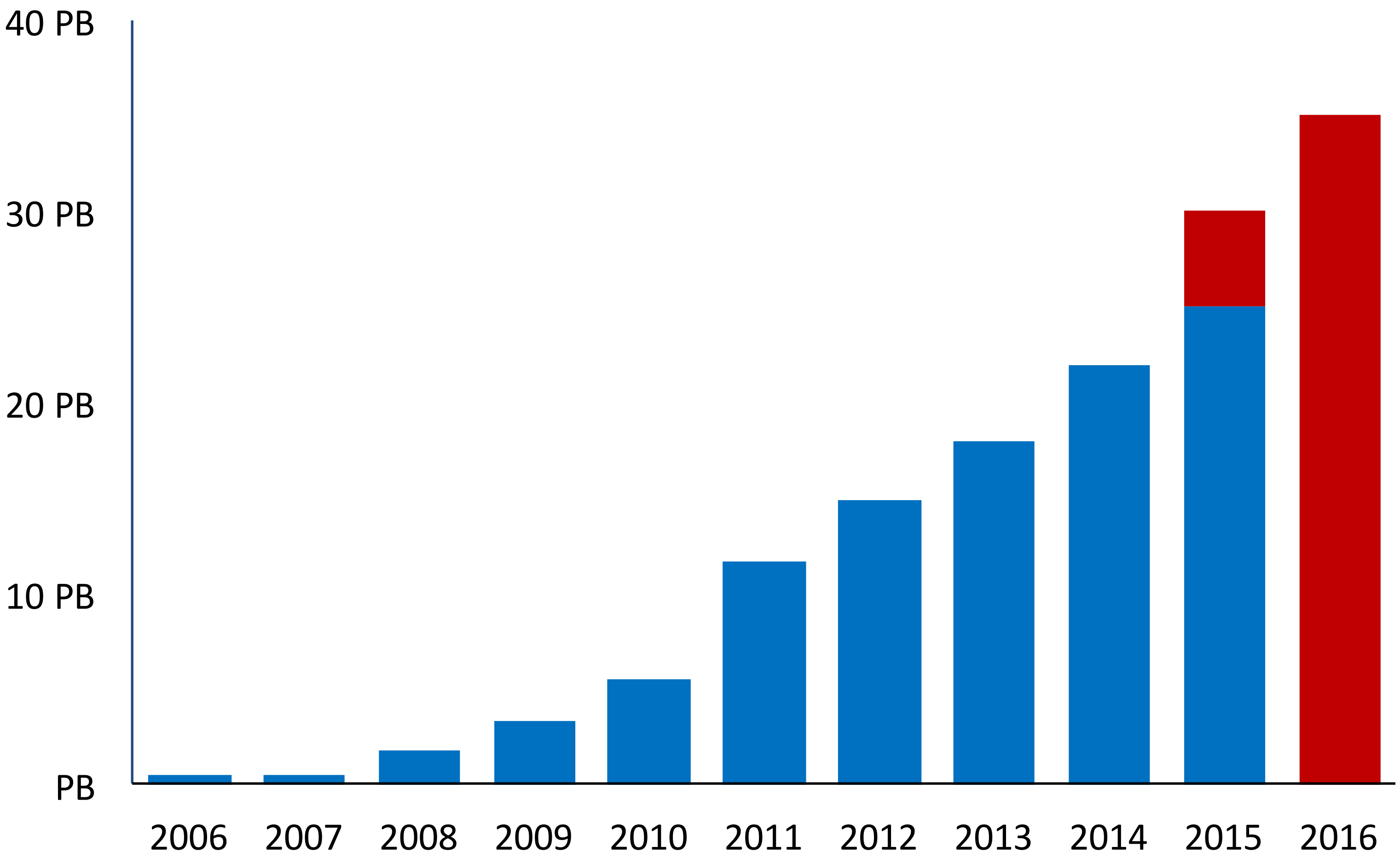
DNA pipelines



wellcome trust sanger institute

Scientific Operations

Data Centre: storage



wellcome trust sanger institute

Scientific Operations

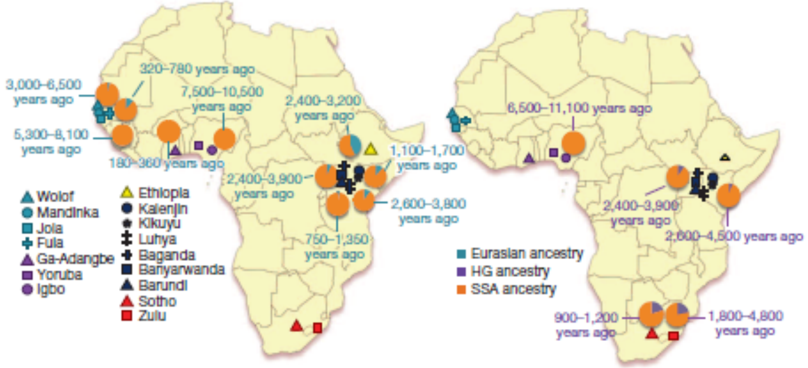
Data Centre

- One of the largest Life Sciences IT infrastructures in the world
- 30 petabytes (10^{15}) storage
- 17,000 CPU cores
- Shared with EBI

What data are we gathering?

The African Genome Variation Project shapes medical genetics in Africa

Deepti Gurdasani[1,2]*, Tommy Carstensen[1,2]*, Fasil Tekola-Ayele[3]*, Luca Pagani[1,4]*, Ioanna Tachmazidou[1]*,
Konstantinos Hatzikotoulas[1], Savita Karthikeyan[1,2], Louise Iles[1,2,5], Martin O. Pollard[1], Ananyo Choudhury[6],
Graham R. S. Ritchie[1,7], Yali Xue[1], Jennifer Asimit[1], Rebecca N. Nsubuga[8], Elizabeth H. Young[1,2], Cristina Pomilla[1,2], Katja Kivinen[1],
Kirk Rockett[9], Anatoli Kamali[8], Ayo P. Doumatey[3], Gershim Asiki[8], Janet Seeley[8], Fatoumatta Sisay-Joof[10], Muminatou Jallow[10],
Stephen Tollman[11,12], Ephrem Mekonnen[13], Rosemary Ekong[14], Tamiru Oljira[15], Neil Bradman[16], Kalifa Bojang[10],
Michele Ramsay[6,17,18], Adebowale Adeyemo[3], Endashaw Bekele[19], Ayesha Motala[20], Shane A. Norris[21], Fraser Pirie[20],
Pontiano Kaleebu[8], Dominic Kwiatkowski[1,9], Chris Tyler-Smith[1]§, Charles Rotimi[3]§, Eleftheria Zeggini[1]§
& Manjinder S. Sandhu[1,2]§

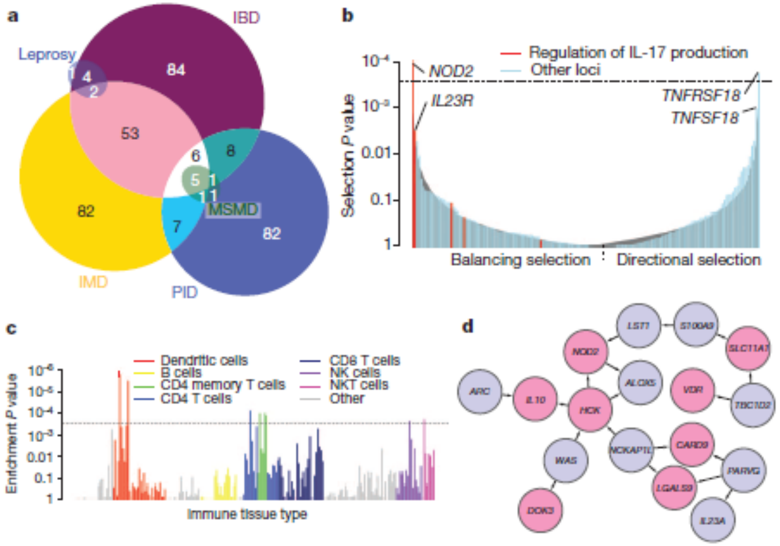


nature

wellcome trust
sanger
institute

What data are we gathering?

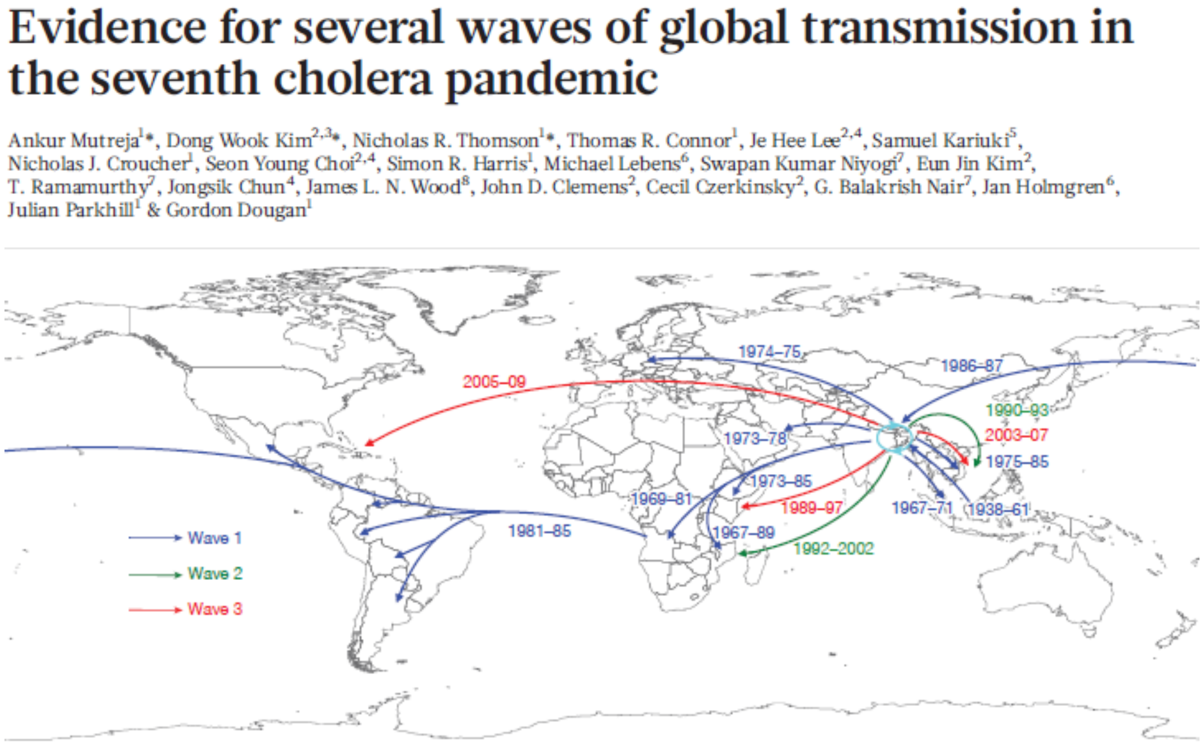
Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease



nature

wellcome trust sanger institute

What data are we gathering?

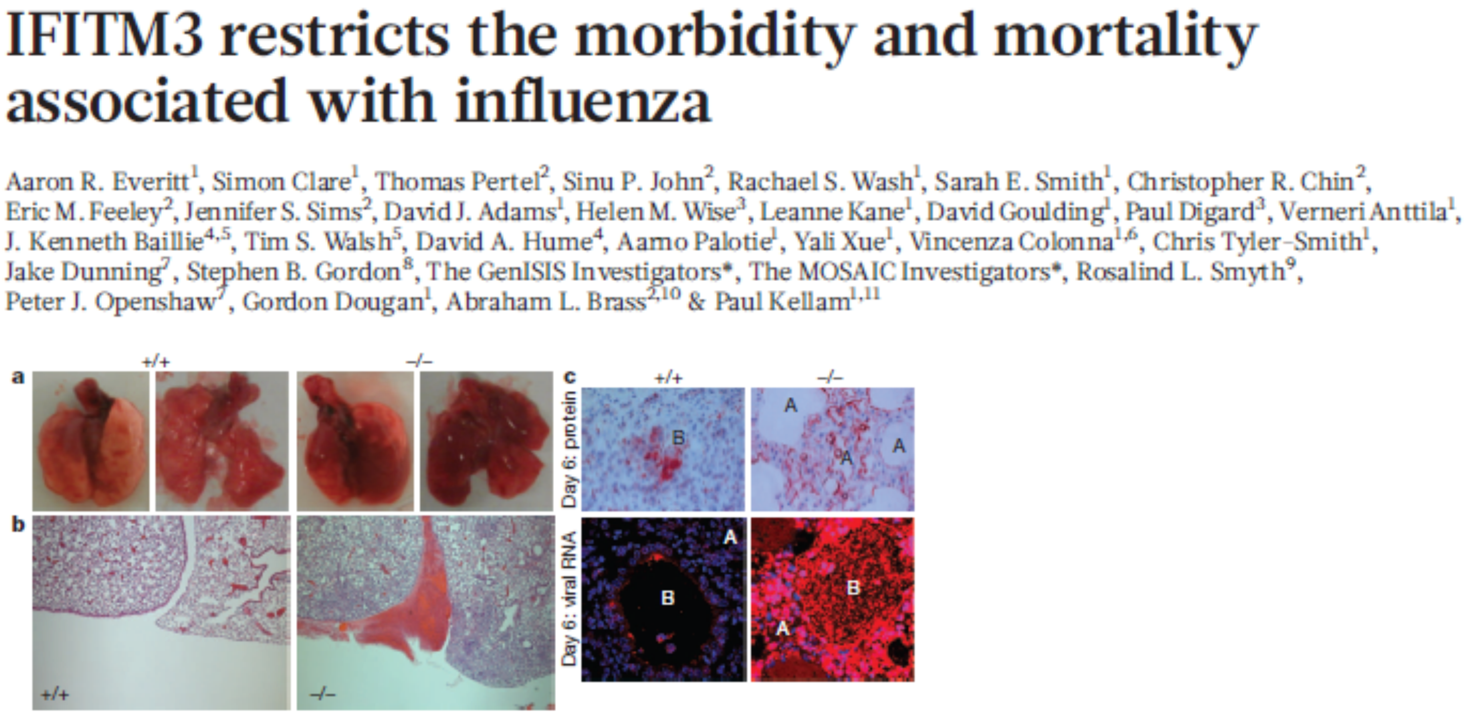


Evidence for several waves of global transmission in the seventh cholera pandemic

Ankur Mutreja[1]*, Dong Wook Kim[2,3]*, Nicholas R. Thomson[1]*, Thomas R. Connor[1], Je Hee Lee[2,4], Samuel Kariuki[5], Nicholas J. Croucher[1], Seon Young Choi[2,4], Simon R. Harris[1], Michael Lebens[6], Swapan Kumar Niyogi[7], Eun Jin Kim[2], T. Ramamurthy[7], Jongsik Chun[4], James L. N. Wood[8], John D. Clemens[2], Cecil Czerkinsky[2], G. Balakrish Nair[7], Jan Holmgren[6], Julian Parkhill[1] & Gordon Dougan[1]

nature

What data are we gathering?



IFITM3 restricts the morbidity and mortality associated with influenza

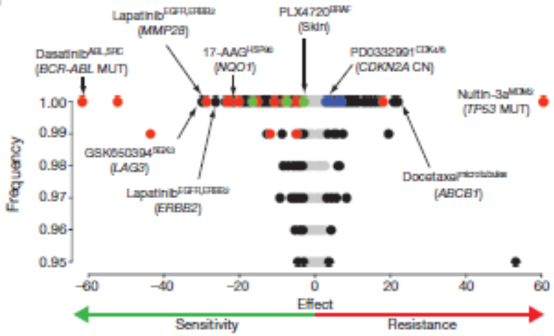
Aaron R. Everitt[1], Simon Clare[1], Thomas Pertel[2], Sinu P. John[2], Rachael S. Wash[1], Sarah E. Smith[1], Christopher R. Chin[2], Eric M. Feeley[2], Jennifer S. Sims[2], David J. Adams[1], Helen M. Wise[3], Leanne Kane[1], David Goulding[1], Paul Digard[3], Verneri Anttila[1], J. Kenneth Baillie[4,5], Tim S. Walsh[5], David A. Hume[4], Aarno Palotie[1], Yali Xue[1], Vincenza Colonna[1,6], Chris Tyler-Smith[1], Jake Dunning[7], Stephen B. Gordon[8], The GenISIS Investigators*, The MOSAIC Investigators*, Rosalind L. Smyth[9], Peter J. Openshaw[7], Gordon Dougan[1], Abraham L. Brass[2,10] & Paul Kellam[1,11]

nature

What data are we gathering?

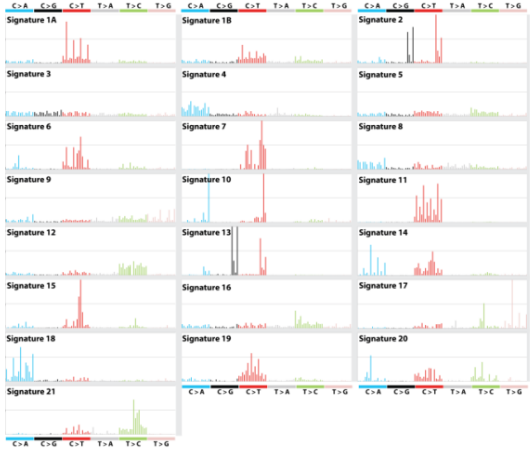
Systematic identification of genomic markers of drug sensitivity in cancer cells

Mathew J. Garnett[1]*, Elena J. Edelman[2]*, Sonja J. Heidorn[1]*, Chris D. Greenman[1]†, Anahita Dastur[2], King Wai Lau[1], Patricia Greninger[2], I. Richard Thompson[1], Xi Luo[2], Jorge Soares[1], Qingsong Liu[3,4], Francesco Iorio[1,5], Didier Surdez[6], Li Chen[2], Randy J. Milano[2], Graham R. Bignell[1], Ah T. Tam[2], Helen Davies[1], Jesse A. Stevenson[2], Syd Barthorpe[1], Stephen R. Lutz[2], Fiona Kogera[1], Karl Lawrence[1], Anne McLaren-Douglas[1], Xeni Mitropoulos[2], Tatiana Mironenko[1], Helen Thi[2], Laura Richardson[1], Wenjun Zhou[3,4], Frances Jewitt[1], Tinghu Zhang[3,4], Patrick O'Brien[1], Jessica L. Boisvert[2], Stacey Price[1], Wooyoung Hur[3,4], Wanjuan Yang[1], Xianming Deng[3,4], Adam Butler[1], Hwan Geun Choi[3,4], Jae Won Chang[3,4], Jose Baselga[2], Ivan Stamenkovic[7], Jeffrey A. Engelman[2], Sreenath V. Sharma[2]†, Olivier Delattre[6], Julio Saez-Rodriguez[5], Nathanael S. Gray[3,4], Jeffrey Settleman[2], P. Andrew Futreal[1], Daniel A. Haber[2,8], Michael R. Stratton[1], Sridhar Ramaswamy[2], Ultan McDermott[1] & Cyril H. Benes[2]



Signatures of mutational processes in human cancer

A list of authors and their affiliations appears at the end of the paper



Sharing our data

	2010	2011	2012	2013	2014
	Visits	**Visits**	**Visits**	**Visits**	**Visits**
Ensembl	3,035,740	3,312,604	3,539,842	3,988,073	4,189,898
COSMIC	234,673	317,157	413,003	467,536	588,281
Decipher	77,987	104,227	113,402	143,449	148,025

Translation

- Centre for Therapeutic Target Validation
- Deciphering Developmental Disorders
- Cancer Cell Line Drug Screen
- Centre for Genomic Pathogen Surveillance

CTTV
Centre for Therapeutic
Target Validation

Centre for
Genomic Pathogen
Surveillance

wellcome trust
sanger
institute

Scientific strategy

By 2020 we will have:

- >100,000 complete or partial human genome sequences from healthy individuals or with disease
- Genome sequences of 30,000 malaria parasites and 20,000 mosquitoes
- Genome sequences of >100,000 bacteria, viruses and parasites
- Complete or partial genome, transcriptome or methylome sequences of 10,000s-100,000s human cells
- Mutant mouse lines for 500-1000 genes
- 1000-2000 iPS cell lines
- 1000-2000 cancer organoids
- 1000s of screens for phenotypic consequences of mutations, drugs and biological challenges

Governing our work

- Balancing the needs of research with the care of research participants

wellcome trust sanger institute

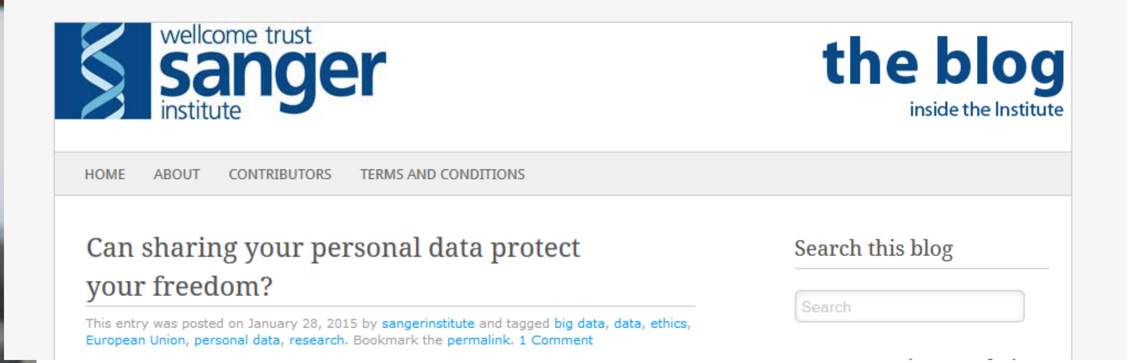
Is this Research or is this Healthcare?

- Deciphering Developmental Disorders
- Sequenced the genomes of 12,600 children and their families
- Diagnoses are returned to families
- Sequences are deposited in DECIPHER

wellcome trust sanger institute

The right to share your own data

- Limiting consent options versus individual freedom



Can sharing your personal data protect your freedom?

This entry was posted on January 28, 2015 by sangerinstitute and tagged big data, data, ethics, European Union, personal data, research. Bookmark the permalink. 1 Comment

Understanding limitations

- Withdrawal of consent for data sharing is not straightforward
- Not the same as consent for participating in research

Image: Wellcome Collection

Ethos of Open Access

- The Sanger Institute founded on the principle that the Human Genome belongs to all

Consent

care.data NHS

- Is opt-out becoming a proxy for consent?

- Research and healthcare are not always distinguishable
- There is a need to recognise individuals' autonomy and rights
- Once shared data cannot be unshared
- Science is a common good and open access is to be commended
- How do we avoid big data using opt-out mechanisms rather than consent?

wellcome trust
sanger
institute